EFFECTS OF HIGH-FREQUENCY CUE REDUCTION ON THE COMPREHENSION OF DISTORTED SPEECH

Paul G. Lacroix and J. Donald Harris

Naval Submarine Medical Research Laboratory, Groton, Connecticut

Forty-five subjects with sharply sloping sensorineural high-frequency hearing losses were examined for comprehension with a tape containing sentences that had been time-compressed (250 words/min), interrupted, (50 msec on-50 msec off), and masked with speech-spectrum noise (+2 dB S/N) in that order. All subjects yielded normal speech reception thresholds, and generally normal scores on the Northwestern University Auditory Test No. 6. Distorted-speech testing was completed at 40 dB SL. Subjects with losses at 2 kHz and above were able to comprehend only 50, 65, and 68% of compressed, interrupted, and noise-masked sentences, respectively. In contrast, subjects with losses at 3 kHz and above performed poorer than normal controls by 11.3, 12.5, and 8 percentage points respectively, while subjects within normal hearing sensitivity at 3 kHz performed as well as controls (maximum drop of 4.6 points with noise-masking). The multiplicative hypothesis was upheld in that the performance of subjects with severe high-frequency deficiencies was much poorer than one would predict on the basis of what is known about performance on these tests with either filtering alone or other distortions alone. These subjects, most especially those with severe frequency deficiencies, outperformed by up to 29 percentage points the mean scores of groups of normals given the same test items and distortion conditions, but listening through filtering that simulated the hearing loss of subjects in this study. It was suggested that hypacusics with high-frequency hearing losses can learn to use residual cues efficiently and that it is not altogether permissible to model sensorineural high-frequency losses with normal-hearing subjects using frequency filtering.

Sher and Owens (1974) reviewed the evidence accumulating from the last 50 years that acoustic cues above 2 kHz are necessary for discriminating isolated words containing certain high-frequency phonemes. There is also mounting evidence that cues above 2 kHz are necessary to extract meaning even from rather highly contextual sentences when the redundant nature of the acoustic, grammatical, lexical, linguistic, and prosodic content of such sentences is reduced by distortion of some sort. It is even likely that the greater the distortion the higher the frequencies required for maximum understanding.

Harris (1960) demonstrated in normal listeners that when different types of distortion, which by themselves are only minimally disruptive, are combined, the cumulative effect on intelligibility will be greater than the sum of the individual effects (the multiplicative hypothesis).

In a preceding paper (Lacroix, Harris, and Randolph, in press) we determined the effects of low-pass (LP) filtering alone, and in combination with time compression, interruption, and noise making, on sentence comprehension by normal listeners. LP filtering of speech, even down to a cutoff of 2

Reprinted from *Journal of Speech and Hearing Disorders*
May 1979, Vol. 44, No. 2

kHz, was quite innocuous, but when only mildly distorted speech was also frequency-filtered, speech comprehension scores dropped by 15-30 percentage points. Evidently, intelligibility can be supported by the rich store of redundancies of all sorts in speech when a single type of cue is reduced; however, if a cue of a different sort is also removed, intelligibility can be significantly affected. For example, the interruption used in the study cited above was arranged at 50 msec on-50 msec off. Here, exactly half the speech is discarded; but although intelligibility was not at all affected, the listener's safety factor was, so to speak, used up. With LP filtering at 2 kHz added, comprehension scores dropped by 22 percentage points, even though LP filtering at 2 kHz by itself had little or no effect on intelligibility.

It remains to be determined whether potentiation occurs when the intrinsic distortions of a hearing loss are combined with extrinsic distortions in the acoustic environment. One may reasonably ask whether a patient with high-tone loss can realistically be modeled as a normal-hearing person listening through an LP filter. As Sher and Owens (1974) said, "The question is whether a normal hearer faced with sudden distortion by low-pass filtering and a hearing-impaired person with 'built-in' distortion from high tone loss corresponding to the 'configuration' imposed by the filter will behave the same with regard to identification of phonemes" (p. 671). Because frequency and related cues are reduced in a similar fashion for both types of listener, one might predict that intelligibility should be equally affected. On the other hand it could be reasoned that there is little similarity between filtering and the distortion imposed by high-frequency hearing loss, since filtering consists only of spectrum limitation, while true hypacusics almost certainly involves additional sources of nonlinearity (Ross, Huntington, Newby, and Dixon, 1965).

In this article, sentence intelligibility in patients with high-frequency hearing losses is reported when the speech was distorted by acceleration, by noise-masking or by interruption. Comparisons are made with normal-hearing subjects who listened to the same distorted sentences. Subgroups of these normal subjects listened through LP filtering which simulated the hearing losses of patients in this article.

METHOD

Subjects

Normals. These were 200 young men, candidates for admission to the U. S. Naval Submarine School, with Hearing Threshold Levels (HTLs) within normal limits from 0.5 to 8 kHz. These subjects were divided into 10 groups 20 men each. Data for nine of these groups were previously reported in a study of combined acoustic distortions (Lacroix et al, in press). The data from the previous study were rescored to correspond to the exact test items listed in the Speech Materials section below. By this method, the previously reported data served as a control against which to compare the performance of our pathological groups. The tenth Normal Group listened to the same recorded speech test that was presented to the pathological subjects.

Pathological Subjects. Forty-five subjects were selected from several hundred patients referred to the Audiology Clinic of this Laboratory for hearing evaluation and grouped as follows:

Group A: 15 patients with pure-tone HTLs $\leqq 20$ dB at 3 kHz and below, but $\geqq 45$ dB at 4 kHz and above;
Group B: 15 patients with pure-tone HTLs $\leqq 20$ dB at 2 kHz and below, but $\geqq 45$ dB at 3 kHz and above; and
Group C: 15 patients with pure-tone HTLs $\leqq 20$ dB at 1 kHz and below, but $\geqq 45$ dB at 2 kHz and above.

Patients were determined by a conventional battery of audiological tests including tympanometry to have nonconductive hearing impairment, much of it noise-induced. Speech Reception Threshold (SRT) in dB for spondees was determined in the experimental ear, and a Discrimination Score (DS) was determined with the Northwestern University Auditory Test No. 6 presented at 40 dB re SL.

All subjects were native speakers of English. Only one ear of each subject was tested.

Speech Materials

A tape of 100 sentences (Harris et al, 1965) was created for this study. In each sentence, three key words lead to a multiple-choice answer. Copies of the master tapes were passed through each of the following three distorting circuits: computer-speeded at 250 words per min, interrupted at 10 ips with 50% duty cycle, and masked at $S/N = 2$ dB by a speech-shaped noise. For all details see Lacroix et al (in press).

Item 1 was discarded, and Items 2-25 used as practice. A final tape was then composed consisting of:

Section A: Items 2-9 time-compressed, serving as practice items for Items 26-50 time-compressed;
Section B: Items 10-17 noise-masked, practice items for following Items 51-75 noise-masked; and
Section C: Items 18-25 interrupted, practice items for following Items 76-100 interrupted.

Procedure

Patients were seated in an IAC 1200 booth with a standard monaural headset. All tapes were played from an Ampex PR10 unit through a speech audiometer calibrated to ANSI standards. Answer sheets were provided, verbal instruction given, and the test started. It should be noted that only 34 patients completed the interruption condition.

Normal subjects listened in 20-man groups. They were seated in a concrete room in the NSMRL sound suite. The chamber is of a tub-within-a-tub construction, separated from outside noises by a 28-inch sandwich of masonry and fibreglass. It is lined with acoustical tile and a cork floor. Each seat was provided with a TDH-39 earphone in an MX-41/AR cushion. Transducers were selected to be matched ±1 dB from 0.1-4 kHz. Noise conditions and transducers for the normal-hearing subjects were in every way comparable to those used for the clinical patients.

All distorted-speech testing was conducted at 40 dB SL.

RESULTS AND DISCUSSION

Table 1 gives the group data (means, standard deviations) for the three groups separately and comparable group data for normal-hearing subjects. Figures 1-3 indicate how closely we were able to construct patient groups with

TABLE 1. DS for hearing-impaired subjects and normal listeners grouped according to hearing loss category and comparable conditions of low-pass filtering.

	Normal—LPF Below 3 kHz*			*Group A (Hearing loss above 3 kHz)*				
	Mean	*SD*	*N*	*Mean*	*SD*	*N*	*Mean Difference*	*P*
DSC (compressed)	78.6	12.5	20	87.7	11.0	15	4.1	0.01
DSC (noise-masked)	83.6	9.9	20	83.7	16.7	15	0.1	n.s.
DSC (interrupted)	94.2	5.6	20	92.9	11.9	15	−1.8	n.s.
DS (NU #6)				97.7	2.7	15		
SRT (dB)				3.8	3.0	15		
	Normal—LPF Below 2 kHz*			*Group B (Hearing loss above 2 kHz)*				
DSC (compressed)	71.4	9.7	20	76.9	12.8	15	4.3	0.001
DSC (noise-masked)	60.2	10.2	20	80.3	12.1	15	15.5	0.001
DSC (interrupted)	89.0	8.2	20	81.3	17.0	9	−4.1	0.001
DS (NU #6)				96.3	4.7	15		
SRT (dB)				5.6	3.1	15		
	Normal—LPF Below 1 kHz*			*Group C (Hearing loss above 1 kHz)*				
DSC (compressed)	29.3	6.7	20	50.5	22.4	15	21.3	0.001
DSC (noise-masked)	45.2	9.7	20	65.2	25.9	15	25.8	0.001
DSC (interrupted)	39.4	10.1	20	68.0	18.3	10	22.8	0.001
DS (NU #6)				76.9	15.2	15		
SRT (dB)				14.2	10.7	15		

Normal—No LPF	*Mean*	*SD*	*N*
DSC (compressed)	88.2	5.5	20
DSC (noise-masked)	88.3	4.9	20
DSC (interrupted)	93.8	6.0	20

*In Tables 2, 3, and 4, and in all figures, some data are included taken from Lacroix et al (1979). Normal-hearing subjects were originally given 50 items both distorted and filtered. These tests were rescored considering only the items heard by the patients of this article, and group data are entered here.

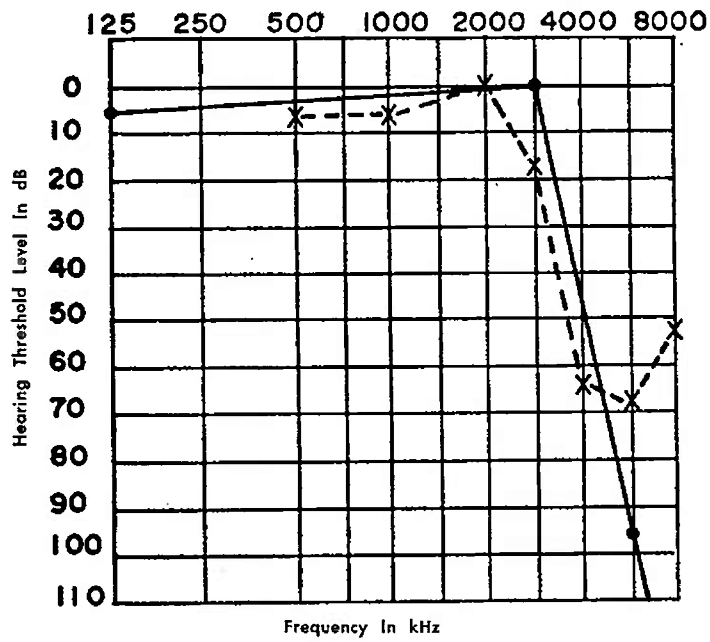

Figure 1. Comparison of mean audiogram of Group A (dashed line: hearing losses above 3 kHz) with the low-pass filter condition imposed on normal listeners in the preceding study (Lacroix et al, 1979) (solid line).

mean audiograms similar to the LP filterings used with the normal-hearing subjects. Figure 4 presents the group means from Table 1, showing the interactions among frequency constraints versus type of distortion for both the normal and the hypacusic ears.

Basic Normalcy of These Patients

As shown in Table 1, all patient groups are within normal limits for SRT and, with the exception of Group C (who had severe losses above 1 kHz), they were also normal for DS by the NU No. 6 test. On the basis of these tests, we have no reason to believe that these patients are in any way unusual in speech-handling capacities except for their frequency-dependent auditory defects.

Effect of Distortions on Normal Listeners

Table 1 shows the mean scores for the normal subjects. These scores of 88.2, 88.3, and 93.8 for compression, noise-making, and interruption, respectively, may be considered norms for this particular tape.

Figure 4 shows that for normals, the effect of filtering is to produce a progressive deterioration as the LP cut-off is rendered more and more severe from 3 to 2 to 1 kHz. Removal of frequencies above 3 kHz has no effect in the interrupted condition.

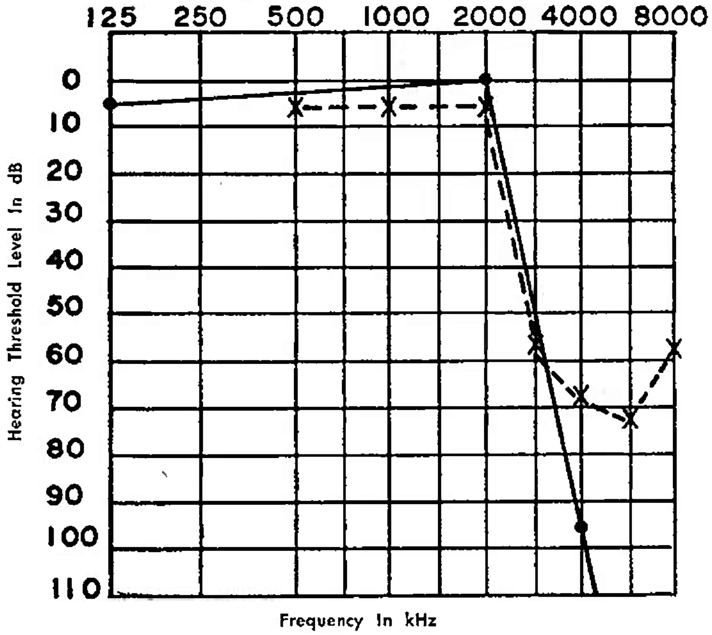

Figure 2. Comparison of mean audiogram of Group B (dashed line: hearing losses above 2 kHz) with the low-pass filter condition imposed on normal listeners in the preceding study (Lacroix et al, 1979) (solid line).

Effects of Distortions on Patients with High-Frequency Hearing Losses

Figure 4 shows regular trends, without inversion, for all types of distortion, as hearing loss becomes more and more severe across Groups A, B, and C respectively. As was the case for normals, the removal of cues above 3 kHz had no effect with interruption.

Among those patients with hearing losses at 2 kHz and above (Groups A and B in Figure 4), mean comprehension scores fall at least 2 standard deviations below the norm. This is true in all cases except for Group B in the noise where the mean score differs from the norm by only 1.5 standard deviations. With the removal of all cues above 1 kHz (Group C), sentence comprehension drops by 23 to 38 percentage points compared with the norm, to a level essentially unusable in everyday situations.

DISCUSSION

Consequences for Assessment of Performance in Everyday Communication Situations

These results furnish another example of the critical nature of the half-octave from 2-3 kHz. For those persons with hearing losses at 3 kHz (Group

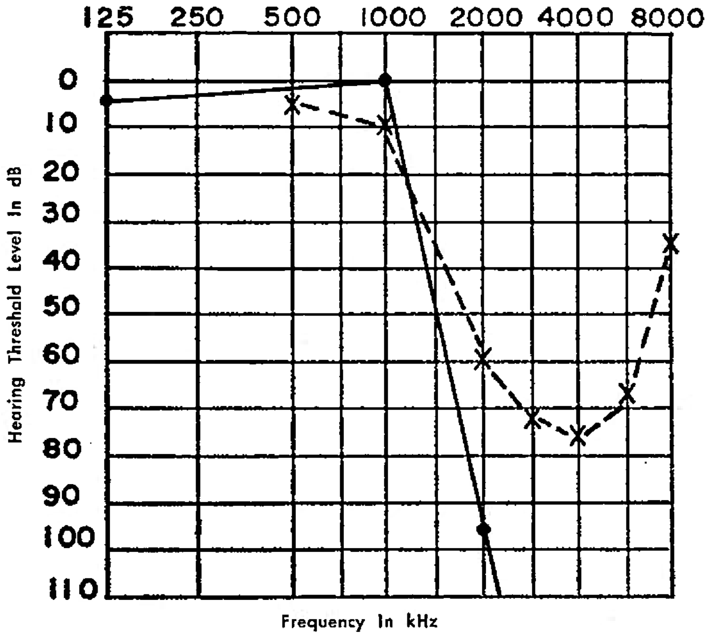

Figure 3. Comparison of mean audiogram of Group C (dashed line: hearing losses above 1 kHz) with the low-pass filter condition imposed on normal listeners in the preceding study (Lacroix et al, 1979) (solid line).

B), a drop in intelligibility of 11.3, 8.0, and 12.5 percentage points occurred for the three distortions. It seems obvious that some account should be taken of the HTL at 3 kHz in formulae designed to predict percentage hearing impairment for everyday situations, or to compute compensation for loss of hearing, since distortions such as these are relatively common. This article does not allow one to say exactly what weight should be attached to HTL at 3 kHz, but our findings do indicate that no consideration need be given to HTL at 4 kHz, because the performance of Group A is well within one standard deviation of normal under all conditions.

The effect of a high-frequency loss in patients is actually a little more severe than appears in Figure 4 because the data are uncorrected for chance. Each sentence is responded to with a four-choice multiple answer, so that chance is 25%. Thus our predictions of the deleterious effects of high-tone losses are on the conservative side.

Interactions Among Effects of Filtering Versus Other Distortions

The experimental design does not permit analysis of variance since obviously no group had two levels of hearing loss. Nevertheless, Figure 4 shows clearly that the effect of spectrum limitation was not the same for each type of distortion. For example, hearing losses above 2 kHz had little effect in the

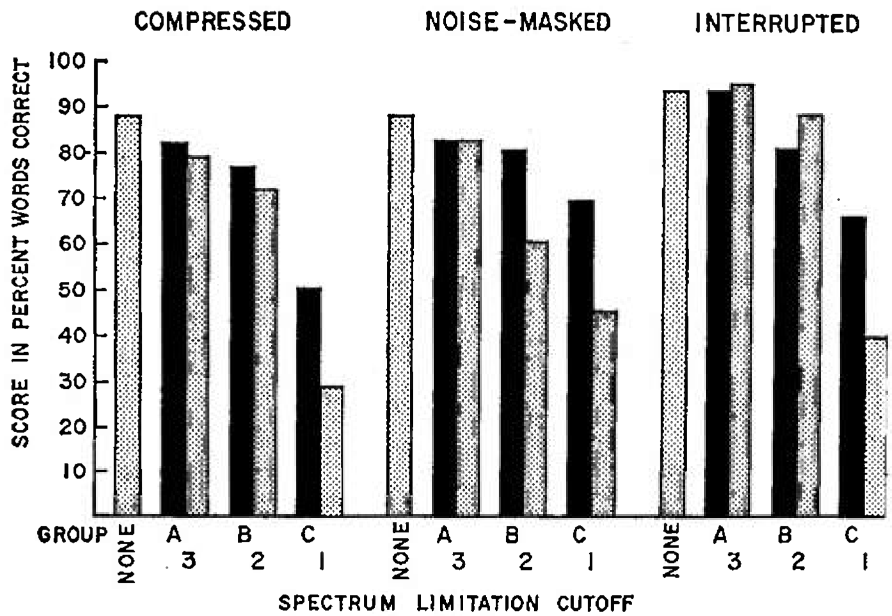

Figure 4. Effect of spectrum limitation on distorted speech comprehension for normal and for hearing-impaired subjects. Hatched: Normals. Solid: Hypacusics. A: Patients with pure tone HTLs $\leqq$ 20 dB at 3 kHz and below, but $\geqq$ 45 dB at 4 kHz and above. B: Patients with pure-tone HTLs $\leqq$ 20 dB at 2 kHz and below, but $\geqq$ 45 dB at 3 kHz and above. C: Patients with pure-tone HTLs $\leqq$ 20 dB at 1 kHz and below, but $\geqq$ 45 dB at 2 kHz and above. None: Normals listening with no frequency filtering. 3: Normals listening through low-pass filtering with cut-off at 3 kHz. 2: Normals listening through low-pass filtering with cut-off at 2 kHz. 1: Normals listening through low-pass filtering with cut-off at 1 kHz. Note that patients usually outperform normal listeners who are given comparable frequency limitation.

TABLE 2. Effect of single distortions on DS for normal listeners (Lacroix et al, 1979).

Condition	*Mean % Correct*	*Difference from Control*
Unfiltered, undistorted (control)	93.9	0
LP at 3 kHz	93.0	0.9
LP at 2 kHz	94.6	−0.7
LP at 1 kHz	86.9	7.0
Average of three distortions	90.1	3.8

interruption condition but a strong effect on noise-masked performance, while hearing loss above 1 kHz had a relatively more serious effect for compression compared with the other two distortions. This suggests that hearing performance is tied to the kind of distortion present.

The Multiplicative Hypothesis

Some light can be shed on the notion that when two types of distortion are combined, each by itself innocuous, the one distortion exerts a potentiating effect on the other and performance drops more than additively. With the present data it was not possible to examine the patients for any distortion without their inherent frequency limitation; however, we may address this question by reference to Tables 2 and 3. Table 2 shows the effects for three

TABLE 3. Comparison of normals and of hypacusics on an average DS from three types of distortion, when groups are matched for spectral limitation.

LP Spectral Limitation in kHz	*Decrease Attributable to Filtering*	*Decrease Attributable to Distortions*		*Predicted Decrease in Intelligibility for Normals*	*Obtained Decrease in Intelligibility for Hypacusics*	*Difference*
3	0.9	+3.8	=	4.7	7.5	−2.8
2	−0.7	+3.8	=	3.1	13.5	−10.4
1	7.0	+3.8	=	10.8	32.7	−21.9

levels of LP filtering in normals and an average for the three distortions given singly to these subjects. By subtracting each of these means in turn from the control condition we obtain an index of the effect, if any, of each separate distortion on intelligibility. Table 3 takes the estimates of individual effects from Table 2 and combines them to predict the additive effect of each combination of filtering and distortion. It is seen that the predicted decreases are minimal, amounting to only 4.7, 3.1, and 10.8% for LP filtering at 3, 2, and 1 kHz respectively. The actual decrements obtained for our groups of patients are all larger and become substantial when low-pass filtering reached 2 kHz. These data indicate that a combination of hearing loss and extrinsic acoustic distortions is likely to be more than additively disruptive to intelligibility.

Differences Between Patients Versus Normals with Simulated Hearing Losses

Figure 4 and Table 1 show significant differences between patients and controls, not always in the same direction. With the most severe high-frequency limitations (Group C; see Table 1), the patients outperform the normals to a very appreciable degree (20-25 points) in all types of distortion. Persons with such hypacusis of long standing may learn to use residual cues more effectively. Whether they use more efficiently the lower-frequency cues for feature extraction, or learn to use more efficiently prosodic or other cues to word prediction/intelligibility cannot be determined from these results. At least it is clear that in experimental situations where a population of persons with high-frequency sensorineural losses is required, it is not altogether permissible to simulate hearing losses by frequency-filtering in normals.

With less severe hearing losses (Group B, see Figure 4 and Table 1), a similar conclusion can be made for noise-masked speech, where the patients outperform controls by 20 points. However, the advantage in the speeded condition is only 5.5 points, while in the interrupted condition the normals have an advantage. The case seems to be that with minor filtering limitations a ceiling effect appears so that the advantage patients may possess, if any, is unnoticeable, although the advantage is clear with more severe filtering.

The results for Group B can be compared with those of a subgroup of 22 patients with sharply sloping audiograms (average drop of 43 dB) from 2-3 kHz, reported by Sher and Owens (1974). On a multiple-choice test where the initial or final consonant only of a CVC word was tested, hypacusics scored 70% correct as against 75.4% for normals given comparable frequency filtering ($P < 0.025$). For patients with more gradual slopes (average drop of 17 dB) there was no deterioration from normal.

Findlay and Denenberg (1977) also compared patients with HTL $\leqq$ 20 dB at 2 kHz but $\geqq$ 40 dB at 4 kHz, with normals given LP filtering at 1.8 kHz (48 dB/octave). On PB lists in babelic noise the patients (DS = 46.7%) outperformed the controls (DS = 37.5%) ($p < 0.05$); this effect was ascribed to the patients having learned to rely on low-frequency cues.

RELATIONS AMONG UNDISTORTED AND DISTORTED SPEECH TESTS

Pearson product-moment rs were calculated among tests for each group separately. Table 4 highlights these data.

The relations among the NU No. 6 test and the distorted-sentence tests is never high. Less than 50%—usually much less—of the variance in the distorted-sentence conditions is common to that obtained with CVC monosyllables in quiet. This lack of prediction from performance in quiet to performance in noise is now common clinical knowledge and needs no further mention.

Correlations among the distortions indicate somewhat higher commonalities, but the interpretation of data for the interrupted condition in Group C is limited by a ceiling effect where eight of the 15 subjects score 100%, and in Groups A and B is limited by the small Ns (10 each). There is reason to think that all three rs for compression versus noise-masking are close to a true value.

These fairly low correlations indicate that a battery of several tests is needed to describe an individual's ability to handle distorted speech; evidently no single test will suffice. It remains to be seen whether three types of distortion

TABLE 4. Pearson product-moment correlation matrix showing relations among speech measures for each hearing-impaired group.

Group	*NU #6 vs Compression*	*NU #6 vs Noise*	*NU #6 vs Interruption*	*Compression vs Interruption*	*Compression vs Noise*	*Interruption vs Noise*
A	−0.17	−0.09	−0.13	0.67	0.63	0.90
B	0.45	0.21	0.60	0.58	0.45	0.50
C	0.74	0.53	0.46	0.67	0.75	0.58

will validly sample the area, either the three types described in this article or some alternative combination.

ACKNOWLEDGMENTS

The authors express grateful appreciation to J. E. Kerivan, Ph.D. and Cathrine Reardon, M.A., for their willing assistance in the clinical testing portion of this study. Requests for reprints should be directed to Paul G. Lacroix, Naval Submarine Medical Research Laboratory, Box 900, Naval Submarine Base New London, Groton, Connecticut 06340.

REFERENCES

FINDLAY, R. C. and DENNENBERG, L. J., Effects of subtle mid-frequency auditory dysfunction upon speech discrimination in noise. *Audiol.*, 16, 252-259 (1977).

HARRIS, J. D., Combinations of distortions in speech. *Arch. Otolaryngol.*, 72, 227-232 (1960).

LACROIX, P. G., HARRIS, J. D. and RANDOLPH, K. J., Multiplicative effects on intelligibility for combined acoustic distortions. *J. Speech Hear. Res.* (in press).

ROSS, M., HUNTINGTON, D. A., NEWBY, H. and DIXON, R. F., Speech discrimination of hearing impaired individuals in noise. *J. Aud. Res.*, 5, 47-72 (1965).

SHER, A. E. and OWENS, E., Consonant confusions associated with hearing loss above 2000 Hz. *J. Speech Hear. Res.*, 17, 669-681 (1974).

Received July 6, 1978.
Accepted December 8, 1978.